EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	26433	(707/1-3,10,102,104.1).CCLS.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 13:58
L2	2375	(709/201).CCLS.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 13:58
L3	0	("(searchnearengine)andindexer").PN.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 13:58
L4	530	(search near engine) and indexer	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:01
L5	19739	1 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:31
L6	326	4 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:16
L7	28382	1 or 2	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:02
L8	187	7 and 6	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:02

L9	10	8 and (anchor near text)	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:27
L10	2387305	9 and descriptive information	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:15
L11	1	9 and descriptive same information	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:15
L12	1	11 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:48
L13	10	9 and indexer	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:28
L14	1	("7181438").PN.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 14:29
L15	8	(("5920859") or ("20050071310") or ("20050060290") or ("20030220912")).PN.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 14:30
L16	6	(("20050071310") or ("20050060290") or ("20030220912")).PN.	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	OFF	2007/03/20 14:31

L17	0	(anchor near text) same (descriptive or reference) near information0 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:32
L18	7	((anchor near text) same ((descriptive or reference) near information))and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:44
L19	0	18 and data near store	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:34
L20	4369	(search near engine) and (indexer or indexing) (anchor near text$1)and crawler	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:48
L21	4369	(search near engine) and (indexer or indexing) (anchor near text$1)and (crawler)	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:48
L22	2977	21 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:52
L23	33249	(reference or descriptive) near information	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:49
L24	10	((reference or descriptive) near information adj2 (data near store))	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:49

L25	2895	10 and 22	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:50
L26	0	24 and 22	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:50
L27	3	anchor near text same (data near store)	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:52
L28	1	27 and @ad<"20031231"	US-PGPUB; USPAT; USOCR; EPO; DERWENT; IBM_TDB	OR	ON	2007/03/20 14:52

IEEE Xplore

Welcome United States Patent and Trademark Office

Search Results   BROWSE   SEARCH   IEEE XPLORE GUIDE

Results for "(('anchor text')<in>metadata)"   e-mail
Your search matched **9** of **1521373** documents.
A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

» Key

IEEE JNL	IEEE Journal or Magazine
IET JNL	IET Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IET CNF	IET Conference Proceeding
IEEE STD	IEEE Standard

Modify Search

(('anchor text')<in>metadata)	Search

☐ Check to search only within this results set

Display Format: ● Citation   ○ Citation & Abstract

[View selected items]   **Select All** **Deselect All**

☐ 1. **Concept-Aware Ranking: Teaching an Old Graph New Moves**
Delong, C.; Mane, S.; Srivastava, J.;
Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE Internation
Dec. 2006 Page(s):80 - 88
Digital Object Identifier 10.1109/ICDMW.2006.49

AbstractPlus | Full Text: PDF(256 KB)   **IEEE CNF**
Rights and Permissions

☐ 2. **A keyword-based semantic prefetching approach in Internet news servici**
Cheng-Zhong Xu; Ibrahim, T.I.;
Knowledge and Data Engineering, IEEE Transactions on
Volume 16, Issue 5, May 2004 Page(s):601 - 611
Digital Object Identifier 10.1109/TKDE.2004.1277820

AbstractPlus | Full Text: PDF(996 KB)   **IEEE JNL**
Rights and Permissions

☐ 3. **Automatic construction of Web directory using hyperlink and anchor text**
Suzuki, Y.; Matsubara, S.; Yoshikawa, M.;
Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-K
Proceedings of 2005 IEEE International Conference on
30 Oct.-1 Nov. 2005 Page(s):614 - 619
Digital Object Identifier 10.1109/NLPKE.2005.1598810

AbstractPlus | Full Text: PDF(4128 KB)   **IEEE CNF**
Rights and Permissions

☐ 4. **Higher-order Web link analysis using multilinear algebra**
Kolda, T.G.; Bader, B.W.; Kenny, J.P.;
Data Mining, Fifth IEEE International Conference on
27-30 Nov. 2005 Page(s):8 pp.
Digital Object Identifier 10.1109/ICDM.2005.77

AbstractPlus | Full Text: PDF(264 KB)   **IEEE CNF**
Rights and Permissions

☐ 5. **Anchor text mining for translation of Web queries**
Wen-Hsiang Lu; Lee-Feng Chien; Hsi-Jian Lee;
Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference or
29 Nov.-2 Dec. 2001 Page(s):401 - 408
Digital Object Identifier 10.1109/ICDM.2001.989545

AbstractPlus | Full Text: PDF(749 KB)   **IEEE CNF**
Rights and Permissions

☐   **6. Web Document Clustering with Multi-view Information Bottleneck**
Yan Gao; Shiwen Gu; Liming Xia; Yaoping Fei;
Computational Intelligence for Modelling, Control and Automation, 2006 and In
Conference on Intelligent Agents, Web Technologies and Internet Commerce,
Conference on
Nov. 2006 Page(s):148 - 148
Digital Object Identifier 10.1109/CIMCA.2006.232

AbstractPlus | Full Text: PDF(204 KB)   **IEEE CNF**
Rights and Permissions

☐   **7. Webpage importance analysis using conditional Markov random walk**
Tie-Yan Liu; Wei-Ying Ma;
Web Intelligence, 2005. Proceedings. The 2005 IEEE/WIC/ACM International (
19-22 Sept. 2005 Page(s):515 - 521
Digital Object Identifier 10.1109/WI.2005.161

AbstractPlus | Full Text: PDF(240 KB)   **IEEE CNF**
Rights and Permissions

☐   **8. Analysis of Topics and Relevant Documents for Navigational Retrieval or**
Oyama, K.; Ishikawa, H.; Aizawa, A.;
Web Information Retrieval and Integration, 2005. WIRI '05. Proceedings. Interr
on Challenges in
08-09 April 2005 Page(s):157 - 163
Digital Object Identifier 10.1109/WIRI.2005.9

AbstractPlus | Full Text: PDF(256 KB)   **IEEE CNF**
Rights and Permissions

☐   **9. Survey of semantic text portion for building Web directory from people's**
Hung, B.Q.; Otsubo, M.; Hijikata, Y.; Nishida, S.;
Active Media Technology, 2005. (AMT 2005). Proceedings of the 2005 Internai
on
19-21 May 2005 Page(s):96 - 101
Digital Object Identifier 10.1109/AMT.2005.1505277

AbstractPlus | Full Text: PDF(385 KB)   **IEEE CNF**
Rights and Permissions

Help   Contact Us   Privacy & :

PORTAL
USPTO

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **anchor text** and **indexer**   Found **447** of **198,991**

Sort results by	relevance	💾 Save results to a Binder	Try an Advanced Search
Display results	expanded form	? Search Tips	Try this search in The ACM Guide
		☐ Open results in a new window	

Results 1 - 20 of 200   Result page: **1** 2 3 4 5 6 7 8 9 10   next
Best 200 shown   Relevance scale ☐ ▭ ▨ ▨ ▨

1 Stanford WebBase components and applications   ▨

Junghoo Cho, Hector Garcia-Molina, Taher Haveliwala, Wang Lam, Andreas Paepcke, Sriram Raghavan, Gary Wesley
May 2006 **ACM Transactions on Internet Technology (TOIT)**, Volume 6 Issue 2
Publisher: ACM Press
Full text available: 📄 pdf(609.18 KB)   Additional Information: full citation, abstract, references, index terms,

We describe the design and performance of WebBase, a tool for Web research. The system includes a highly customizable crawler, a repository for collected Web pages, an indexer for both text and link-related page features, and a high-speed content distribution facility. The distribution module enables researchers world-wide to retrieve pages from WebBase, and stream them across the Internet at high speed. The advantage for the researchers is that they need not all crawl the Web before beginning t ...

Keywords: WebBase Web crawler, distribution, hyperlink indexing, site crawling

2 Topical locality in the Web   ▨

Brian D. Davison
July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '00**
Publisher: ACM Press
Full text available: 📄 pdf(771.77 KB)   Additional Information: full citation, abstract, references, citings, index terms

Most web pages are linked to others with related content. This idea, combined with another that says that *text in, and possibly around, HTML anchors describe the pages to which they point,* is the foundation for a usable World-Wide Web. In this paper, we examine to what extent these ideas hold by empirically testing whether topical locality mirrors spatial locality of pages on the Web. In particular, we find that the likelihood of linked pages having similar textual content to be ...

3 Finding links   ▨

John Tebbutt
May 1998 **Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems HYPERTEXT '98**

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **anchor**
text and **indexer** and **reference** or **descriptive** and **information**

Found **130,664** of **198,991**

Sort results
by | relevance ▼ |

Display
results | expanded form ▼ |

Save results to a Binder
Search Tips
☐ Open results in a new
window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200   Result page: **1** 2 3 4 5 6 7 8 9 10   next
Best 200 shown   Relevance scale ☐ ▭ ▦ ▦ ▦

1   Topical locality in the Web
Brian D. Davison
July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on
Research and development in information retrieval SIGIR '00**
Publisher: ACM Press

Full text available: ▦ pdf(771.77 KB)   Additional Information: full citation, abstract, references, citings, index
terms

Most web pages are linked to others with related content. This idea, combined with
another that says that *text in, and possibly around, HTML anchors describe the pages to
which they point*, is the foundation for a usable World-Wide Web. In this paper, we
examine to what extent these ideas hold by empirically testing whether topical locality
mirrors spatial locality of pages on the Web. In particular, we find that the likelihood of
linked pages having similar textual content to be ...

2   Shape-based retrieval and analysis of 3D models
Thomas Funkhouser, Michael Kazhdan
August 2004 **ACM SIGGRAPH 2004 Course Notes SIGGRAPH '04**
Publisher: ACM Press

Full text available: ▦ pdf(12.56 MB)   Additional Information: full citation, abstract

Large repositories of 3D data are rapidly becoming available in several fields, including
mechanical CAD, molecular biology, and computer graphics. As the number of 3D models
grows, there is an increasing need for computer algorithms to help people find the
interesting ones and discover relationships between them. Unfortunately, traditional text-
based search techniques are not always effective for 3D models, especially when queries
are geometric in nature (e.g., find me objects that fit into thi ...

3   Searching the Web
Arvind Arasu, Junghoo Cho, Hector Garcia-Molina, Andreas Paepcke, Sriram Raghavan
August 2001 **ACM Transactions on Internet Technology (TOIT)**, Volume 1 Issue 1
Publisher: ACM Press

Full text available: ▦ pdf(319.98 KB)   Additional Information: full citation, abstract, references, citings, index
terms, review

We offer an overview of current Web search engine design. After introducing a generic
search engine architecture, we examine each engine component in turn. We cover
crawling, local Web page storage, indexing, and the use of link analysis for boosting

P@RTAL
USPTO

Search: ⊙ The ACM Digital Library   ○ The Guide

"anchor text" and indexer and (reference or descriptive) and in

THE ACM DIGITAL LIBRARY

Feedback   Report a problem   Satisfaction survey

Terms used **anchor text** and **indexer** and **reference** or **descriptive** and **information** and **data store**   Found **95,933** of **198,991**

Sort results by   relevance ▾

Display results   expanded form ▾

🔖 Save results to a Binder
❓ Search Tips
☐ Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200   Result page: **1** 2 3 4 5 6 7 8 9 10   next

Best 200 shown   Relevance scale ☐ ▨ ▨ ▨ ▨

1 Information storage and retrieval: a survey and functional description   ▨

Jack Minker
September 1977 **ACM SIGIR Forum**, Volume 12 Issue 2
Publisher: ACM Press
Full text available: 📄 pdf(5.14 MB)   Additional Information: full citation, abstract, references, citings

Information Storage and Retrieval (IS&R) encompasses a broad scope of topics ranging from basic techniques for accessing data to sophisticated approaches for the analysis of natural language text and the deduction of information. Within the field, three general areas of investigation can be distinguished not only by their subject matter but also by the types of individuals presently interested in them:(1) Document retrieval,(2) Generalized data management, and(3) Question-answering.A functional ...

Keywords: automatic indexing, data management, data structures, deductive search, information retrieval, natural language, problem solving, question-answering, relational data systems, theorem proving

2 Topical locality in the Web   ▨

Brian D. Davison
July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '00**
Publisher: ACM Press
Full text available: 📄 pdf(771.77 KB)   Additional Information: full citation, abstract, references, citings, index terms

Most web pages are linked to others with related content. This idea, combined with another that says that *text in, and possibly around, HTML anchors describe the pages to which they point,* is the foundation for a usable World-Wide Web. In this paper, we examine to what extent these ideas hold by empirically testing whether topical locality mirrors spatial locality of pages on the Web. In particular, we find that the likelihood of linked pages having similar textual content to be ...

3 Representing internet streaming media metadata using MPEG-7 multimedia description schemes   ▨

Eric Rehm

November 2000 **Proceedings of the 2000 ACM workshops on Multimedia MULTIMEDIA '00**
Publisher: ACM Press
Full text available: pdf(603.51 KB)   Additional Information: full citation, abstract, references, index terms

> Singingfish.com uses MPEG-7 description schemes to model the metadata characteristics of Internet streaming media. Further it has used MPEG-7 description schemes as the basis for a number of internal experiments, processes, as well as interchange with commercial content partners. This paper describes our selection of description schemes from the MPEG-7 Multimedia Description Schemes (MDS) Working Drafts, their ability (or lack thereof) to represent Internet streaming media metadata, and our e ...

> **Keywords**: MPEG-7, multimedia, search engine, streaming media

4 Hypertext versioning: Automatic generation of hypertext system repositories: a model driven approach
E. James Whitehead, Guozheng Ge, Kai Pan
August 2004 **Proceedings of the fifteenth ACM conference on Hypertext and hypermedia HYPERTEXT '04**
Publisher: ACM Press
Full text available: pdf(593.73 KB)   Additional Information: full citation, abstract, references, citings, index terms

> In this paper, we present a model-driven methodology and toolset for automatic generation of hypertext system repositories. Our code generator, called Bamboo, is based on a Containment Modeling Framework (CMF) that uniformly describes data models for hypertext systems. CMF employs a lightweight modeling approach in which entities (system abstractions) and containment relationships are used to model hypertext system repositories. Given a description of a system repository data model using CMF, as ...

> **Keywords**: automatic code generation, containment modeling framework, hypertext data models, model-driven development, open hyperbase

5 Stanford WebBase components and applications
Junghoo Cho, Hector Garcia-Molina, Taher Haveliwala, Wang Lam, Andreas Paepcke, Sriram Raghavan, Gary Wesley
May 2006 **ACM Transactions on Internet Technology (TOIT)**, Volume 6 Issue 2
Publisher: ACM Press
Full text available: pdf(609.18 KB)   Additional Information: full citation, abstract, references, index terms

> We describe the design and performance of WebBase, a tool for Web research. The system includes a highly customizable crawler, a repository for collected Web pages, an indexer for both text and link-related page features, and a high-speed content distribution facility. The distribution module enables researchers world-wide to retrieve pages from WebBase, and stream them across the Internet at high speed. The advantage for the researchers is that they need not all crawl the Web before beginning t ...

> **Keywords**: WebBase Web crawler, distribution, hyperlink indexing, site crawling

6 Shape-based retrieval and analysis of 3D models
Thomas Funkhouser, Michael Kazhdan
August 2004 **ACM SIGGRAPH 2004 Course Notes SIGGRAPH '04**
Publisher: ACM Press
Full text available: pdf(12.56 MB)   Additional Information: full citation, abstract

Large repositories of 3D data are rapidly becoming available in several fields, including mechanical CAD, molecular biology, and computer graphics. As the number of 3D models grows, there is an increasing need for computer algorithms to help people find the interesting ones and discover relationships between them. Unfortunately, traditional text-based search techniques are not always effective for 3D models, especially when queries are geometric in nature (e.g., find me objects that fit into thi ...

7 Searching the Web

Arvind Arasu, Junghoo Cho, Hector Garcia-Molina, Andreas Paepcke, Sriram Raghavan

August 2001 **ACM Transactions on Internet Technology (TOIT)**, Volume 1 Issue 1

Publisher: ACM Press

Full text available: pdf(319.98 KB)   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>, <u>review</u>

We offer an overview of current Web search engine design. After introducing a generic search engine architecture, we examine each engine component in turn. We cover crawling, local Web page storage, indexing, and the use of link analysis for boosting search performance. The most common design and implementation techniques for each of these components are presented. For this presentation we draw from the literature and from our own experimental search engine testbed. Emphasis is on introduci ...

Keywords: HITS, PageRank, authorities, crawling, indexing, information retrieval, link analysis, search engine

8 Establishing the semantic web 1: SemTag and seeker: bootstrapping the semantic web via automated semantic annotation

Stephen Dill, Nadav Eiron, David Gibson, Daniel Gruhl, R. Guha, Anant Jhingran, Tapas Kanungo, Sridhar Rajagopalan, Andrew Tomkins, John A. Tomlin, Jason Y. Zien

May 2003 **Proceedings of the 12th international conference on World Wide Web WWW '03**

Publisher: ACM Press

Full text available: pdf(178.36 KB)   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

This paper describes Seeker, a platform for large-scale text analytics, and SemTag, an application written on the platform to perform automated semantic tagging of large corpora. We apply SemTag to a collection of approximately 264 million web pages, and generate approximately 434 million automatically disambiguated semantic tags, published to the web as a label bureau providing metadata regarding the 434 million annotations. To our knowledge, this is the largest scale semantic tagging effort to ...

Keywords: automated semantic tagging, data mining, information retrieval, large text datasets, text analytics

9 Automatic subject indexing using an associative neural network

Yi-Ming Chung, William M. Pottenger, Bruce R. Schatz

May 1998 **Proceedings of the third ACM conference on Digital libraries DL '98**

Publisher: ACM Press

Full text available: pdf(1.26 MB)   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

10 Database techniques for the World-Wide Web: a survey

Daniela Florescu, Alon Levy, Alberto Mendelzon

September 1998 **ACM SIGMOD Record**, Volume 27 Issue 3
Publisher: ACM Press
Full text available: pdf(1.79 MB)   Additional Information: full citation, citings, index terms

---

11 ANNOD: a navigator of natural-language organized (textual) data

Robert E. Williamson
June 1985 **Proceedings of the 8th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '85**
Publisher: ACM Press
Full text available: pdf(830.49 KB)   Additional Information: full citation, abstract, references

ANNOD is the name of a system developed at the National Library of Medicine (NLM), which implements a set of linguistic and empirical techniques that permit retrieval of natural language information in response to natural language queries. The system is based on Dr. Gerard Salton's SMART [1] document retrieval system and is presently implemented on a mini-computer as part of an Interactive TExt Management System, ITEMS.[2] Actual experience with retrieval of information from NLM's Hepatitis ...

---

12 Poster papers: Discovering informative content blocks from Web documents

Shian-Hua Lin, Jan-Ming Ho
July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining KDD '02**
Publisher: ACM Press
Full text available: pdf(692.97 KB)   Additional Information: full citation, abstract, references, citings, index terms

In this paper, we propose a new approach to discover informative contents from a set of tabular documents (or Web pages) of a Web site. Our system, InfoDiscoverer, first partitions a page into several content blocks according to HTML tag <TABLE> in a Web page. Based on the occurrence of the features (terms) in the set of pages, it calculates entropy value of each feature. According to the entropy value of each feature in a content block, the entropy value of the block is defined. By analyz ...

Keywords: entropy, information extraction, information retrieval, informative content discovery

---

13 Learning classifiers: Liveclassifier: creating hierarchical text classifiers through web corpora

Chien-Chung Huang, Shui-Lung Chuang, Lee-Feng Chien
May 2004 **Proceedings of the 13th international conference on World Wide Web WWW '04**
Publisher: ACM Press
Full text available: pdf(244.58 KB)   Additional Information: full citation, abstract, references, citings, index terms

Many Web information services utilize techniques of information extraction(IE) to collect important facts from the Web. To create more advanced services, one possible method is to discover thematic information from the collected facts through text classification. However, most conventional text classification techniques rely on manual-labelled corpora and are thus ill-suited to cooperate with Web information services with open domains. In this work, we present a system named LiveClassifier that ...

Keywords: text classification, topic hierarchy, web mining

14 Time and space: Time period directories: a metadata infrastructure for placing events in temporal and geographic context

Vivien Petras, Ray R. Larson, Michael Buckland

June 2006 **Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries JCDL '06**

Publisher: ACM Press

Full text available: pdf(418.63 KB)   Additional Information: full citation, abstract, references, index terms

> Metadata is ordinarily used to describe documents, but it can also constitute a form of infrastructure for access to networked resources and for traversal of those resources. One problematic area for access to digital library resources has been the search for time periods or events. If there is a capability to search for time, it is usually a date search - a standardized and precise form but unfortunately rarely used in common chronological expressions. For example, a user interested in the "Vie ...
>
> **Keywords:** metadata infrastructure, time period directories

15 Demo sessions: Demo group A: Query and search: SIREN: a similarity retrieval engine for complex data

Maria Camila N. Barioni, Humberto Razente, Agma Traina, Caetano Traina

September 2006 **Proceedings of the 32nd international conference on Very large data bases - Volume 32 VLDB'2006**

Publisher: VLDB Endowment

Full text available: pdf(748.60 KB)   Additional Information: full citation, abstract, references, index terms

> This paper presents a similarity retrieval engine - SIREN-that allows posing similarity queries in a relational DBMS using an extended syntax that adds the support for such type of queries in the SQL language. It discusses the main architecture of SIREN, describes some key features and provides a description of the demo.

16 PYTHIA-II: a knowledge/database system for managing performance data and recommending scientific software

Elias N. Houstis, Ann C. Catlin, John R. Rice, Vassilios S. Verykios, Naren Ramakrishnan, Catherine E. Houstis

June 2000 **ACM Transactions on Mathematical Software (TOMS)**, Volume 26 Issue 2

Publisher: ACM Press

Full text available: pdf(796.18 KB)   Additional Information: full citation, abstract, references, citings, index terms

> Often scientists need to locate appropriate software for their problems and then select from among many alternatives. We have previously proposed an approach for dealing with this task by processing performance data of the targeted software. This approach has been tested using a customized implementation referred to as PYTHIA. This experience made us realize the complexity of the algorithmic discovery of knowledge from performance data and of the management of these data together with the d ...
>
> **Keywords:** data mining, inductive logic programming, knowledge discovery in databases, knowledge-based systems, performance evaluation, recommender systems, scientific software

17 Reference model for DBMS standardization

March 1986 **ACM SIGMOD Record**, Volume 15 Issue 1

Publisher: ACM Press

Full text available: pdf(2.62 MB)   Additional Information: full citation, abstract, citings, index terms

This report proposes a Reference Model (RM) for database management system (DBMS) standardization. A Reference Model is a conceptual framework whose purpose is to divide standardization work into manageable pieces and to show at a general level how these pieces are related with each other. The proposed RM comprises a Data Mapping Control System (DMCS) that retrieves and stores application data, application schemas, and data dictionary schemas. This DMCS is bounded by two interfaces: the Data Lan ...

18 Web search 2: A study of relevance propagation for web search

Tao Qin, Tie-Yan Liu, Xu-Dong Zhang, Zheng Chen, Wei-Ying Ma
August 2005 **Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '05**
Publisher: ACM Press

Full text available: 📄 pdf(391.03 KB)   Additional Information: full citation, abstract, references, citings, index terms

Different from traditional information retrieval, both content and structure are critical to the success of Web information retrieval. In recent years, many relevance propagation techniques have been proposed to propagate content information between web pages through web structure to improve the performance of web search. In this paper, we first propose a generic relevance propagation framework, and then provide a comparison study on the effectiveness and efficiency of various representative pro ...

Keywords: hyperlink based score propagation, hyperlink based term propagation, relevance propagation, sitemap based score propagation, sitemap based term propagation

19 Charles W. Bachman interview: September 25-26, 2004; Tucson, Arizona

Thomas Haigh
January 2006 **ACM Oral History interviews**
Publisher: ACM Press
Full text available: 📄 pdf(761.66 KB)   Additional Information: full citation, abstract

Charles W. Bachman reviews his career. Born during 1924 in Kansas, Bachman attended high-school in East Lansing, Michigan before joining the Army Anti Aircraft Artillery Corp, with which he spent two years in the Southwest Pacific Theater, during World War II. After his discharge from the military, Bachman earned a B.Sc. in Mechanical Engineering in 1948, followed immediately by an M.Sc. in the same discipline, from the University of Pennsylvania. On graduation, he went to work for Do ...

20 Ranking and estimation: Estimating corpus size via queries

Andrei Broder, Marcus Fontura, Vanja Josifovski, Ravi Kumar, Rajeev Motwani, Shubha Nabar, Rina Panigrahy, Andrew Tomkins, Ying Xu
November 2006 **Proceedings of the 15th ACM international conference on Information and knowledge management CIKM '06**
Publisher: ACM Press
Full text available: 📄 pdf(285.61 KB)   Additional Information: full citation, abstract, references, index terms

We consider the problem of estimating the size of a collection of documents using only a standard query interface. Our main idea is to construct an unbiased and low-variance estimator that can closely approximate the size of any set of documents defined by certain conditions, including that each document in the set must match at least one query from a uniformly sampleable query pool of known size, fixed in advance.Using this basic estimator, we propose two approaches to estimating corpus size. T ...

Keywords: corpus size, estimator, random sampling

Results 1 - 20 of 200   Result page: **1**   2   3   4   5   6   7   8   9   10   next

Useful downloads: Adobe Acrobat   QuickTime   Windows Media Player   Real Player